NOTE ON A CASE-STUDY IN BOX-JENKINS SEASONAL FORECASTING OF TIME SERIES

BY

STEFFEN L. LAURITZEN

TECHNICAL REPORT NO. 16
APRIL 1974

PREPARED UNDER CONTRACT
N00014-67-A-0112-0030 (NR-042-034)
FOR THE OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

NOTE ON A CASE-STUDY IN BOX-JENKINS SEASONAL FORECASTING OF TIME SERIES

by

STEFFEN L. LAURITZEN

STANFORD UNIVERSITY

TECHNICAL REPORT NO. 16

APRIL 1974

PREPARED UNDER THE AUSPICES OF

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030

T. W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Recently, the Journal of the Royal Statistical Society brought a lively discussion of the Box-Jenkins procedure for forecasting seasonal time series. This took place in two papers, one by Chatfield and Prothero (1973) and one by Box and Jenkins (1973). The discussion was centered around a particular case study of some sales figures.

Box and Jenkins refer to several other case-studies in their reply, including a paper on telephone installations and removals by Thompson and Tiao (1971). The original data in this analysis are not in the paper, but can be found in the preceding technical report from University of Wisconsin [Thompson and Tiao (1970)]. The present author had some time ago an opportunity to look at these data and the findings might be of some general interest in view of the recent discussion.

The data consist of the monthly totals of telephone installations and removals made by the Wisconsin Telephone Company from January 1951 through October 1966. The objective of the analysis is to obtain good forecasts for the two series. The authors consider the logarithms of the two series, respectively denoted by X_t and Y_t. Via the procedure described in the book by Box and Jenkins (1970) ARIMA models are obtained for $Z_t = X_t - Y_t$ and Y_t. The parameters are estimated and forecasts made according to the estimated models. The forecasts of both series from November 1966 through October 1968 are compared to actual observations with quite good results. Following the standard notation the models for the Z_t and Y_t series were as follows:

$$(1 - B^{12})(1 - \phi_1 B) Z_t = (1 - \theta_{12} B^{12})(1 - \theta_1 B) a_t , \qquad (1)$$

where a_t are uncorrelated random disturbances and

$$(1 - \phi_3 B^3)\ (1 - \phi_{12} B^{12})\ Y_t = (1 - \delta_9\ B^9 - \delta_{12}\ B^{12} - \delta_{13}\ B^{13}) b_t, \qquad (2)$$

The model (2) seems to be quite complicated, and (also suggested by Thompson and Tiao) calendar irregularities might be the reason. If one considers the number of installations or removals in any given month to be roughly proportional to the number of weekdays in that month, the effect of the calendar irregularities would be present in the Y_t-series but not in the Z_t-series as this involves only the ratio between the number of installations and removals. So as a simple method of forecasting the number of removals in any given month one might suggest that the average daily number of removals made in that month would increase by a fixed percentage over the corresponding number in the same month of the preceding year. That is, if we denote the number of weekdays in month t by d_t, one could let the forecast of $Y_{T+12k+j}$, $1 \leqq j \leqq 12$, from $Y_1, \ldots, Y_T$ be

$$\hat{Y}_{T+12k+j} = Y_{T+j-12} - \log d_{T+j-12} + \log d_{T+12k+j} + (k+1)\ \theta\ . \qquad (3)$$

Averaging over $(Y_t - \log d_t) = (Y_{t-12} - \log d_{t-12})$ suggests that $\theta = 0.049$ would be a good value for θ, corresponding to a 5% yearly increase.

The author tried to do exactly as in the paper by Thompson and Tiao apart from substituting the "naive" method (3) for the forecasts obtained from the model (2). That is Z_t was forecasted from the model (1) and Y_t by (3) and then $X_t = Y_t + Z_t$ forecasted as

$$\hat{X}_{T+i} = \hat{Y}_{T+i} + \hat{Z}_{T+i}\ . \qquad (4)$$

Tables 1 and 2 show the percentage error of forecasts for the total number of installations and removals, respectively, obtained by the two

methods for the 24 months from November 1966 through October 1968. These data are plotted on Figures 1 and 2.

Table 1

Percentage Error in Forecasts of the Number of Telephone Removals

Period	Nov. 1966	Dec.	Jan. 1967	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.
% error by original method	+6	+6	-5	+3	+15	+12	-4	+1	+3	+7	+4	-7
% error by naive method	0	-1	0	+1	+6	+5	0	+1	+2	+8	0	-2

Period	Nov.	Dec.	Jan. 1968	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.
% error by original method	+8	+6	-10	-5	+14	+14	-12	+7	-12	+9	+7	0
% error by naive method	+8	-10	-5	-7	+2	+17	-8	+8	-4	+5	+3	+9

Table 2

Percentage Error in Forecasts of the Number of Telephone Installations

Period	Nov. 1966	Dec.	Jan. 1967	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.
% error by original method	+9	+5	+2	+5	+9	+10	+2	-2	+7	+7	+6	-6
% error by naive method	+3	-2	+7	+3	0	+3	+6	-2	+6	+8	+2	-1

Period	Nov. 1967	Dec.	Jan. 1968	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.
% error by original method	+8	+1	-3	-2	+14	+26	-10	+14	-11	0	+6	-6
% error by naive method	+8	-15	+2	-4	+2	+29	-6	+15	-3	-4	+2	+3

It is surprising but obvious that the "naive" method seems to give much better results for the forecasts of the removals and it does not do worse for the forecasts of the number of installations. None of the methods could forecast the strike in April 1968.

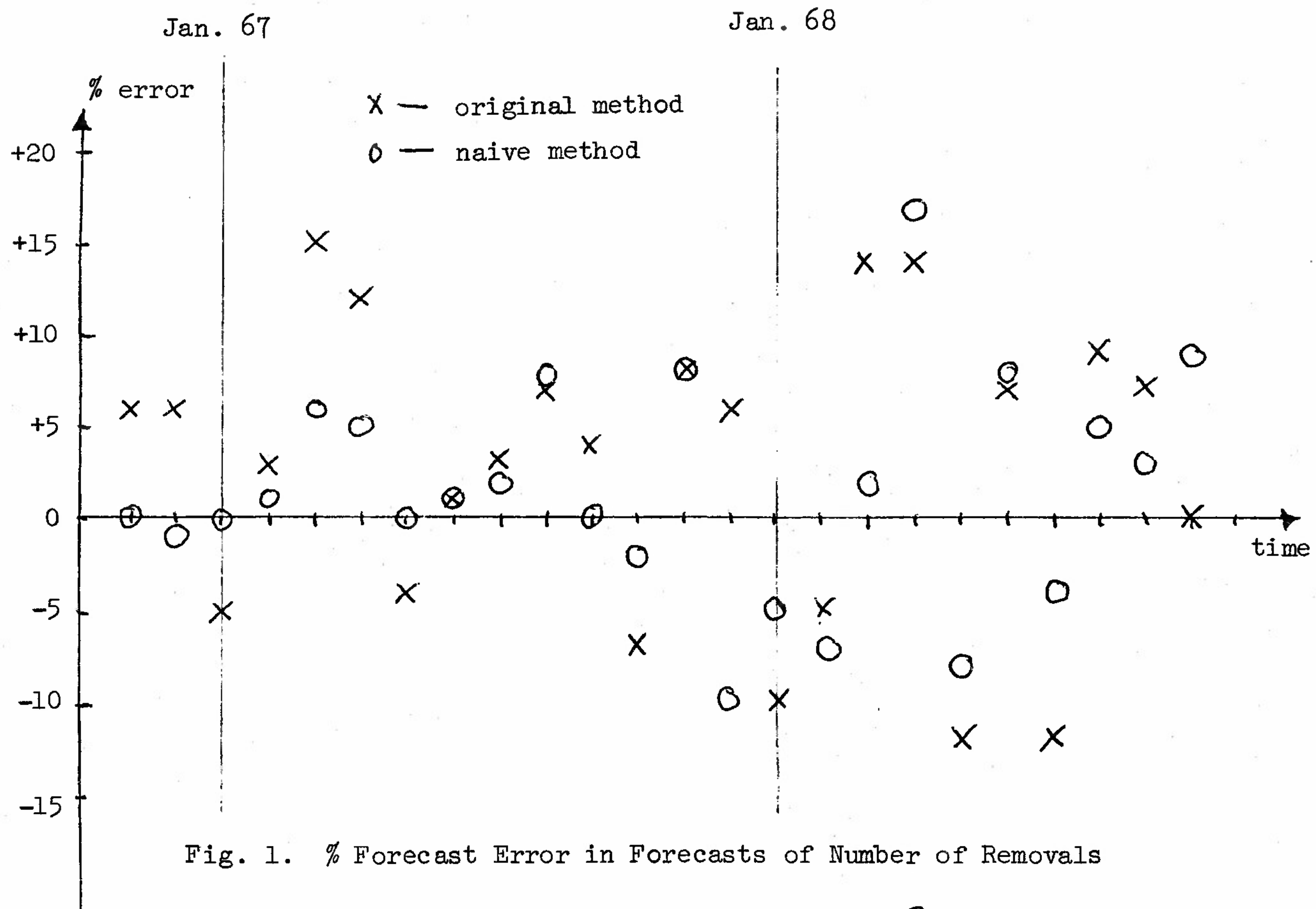

Fig. 1. % Forecast Error in Forecasts of Number of Removals

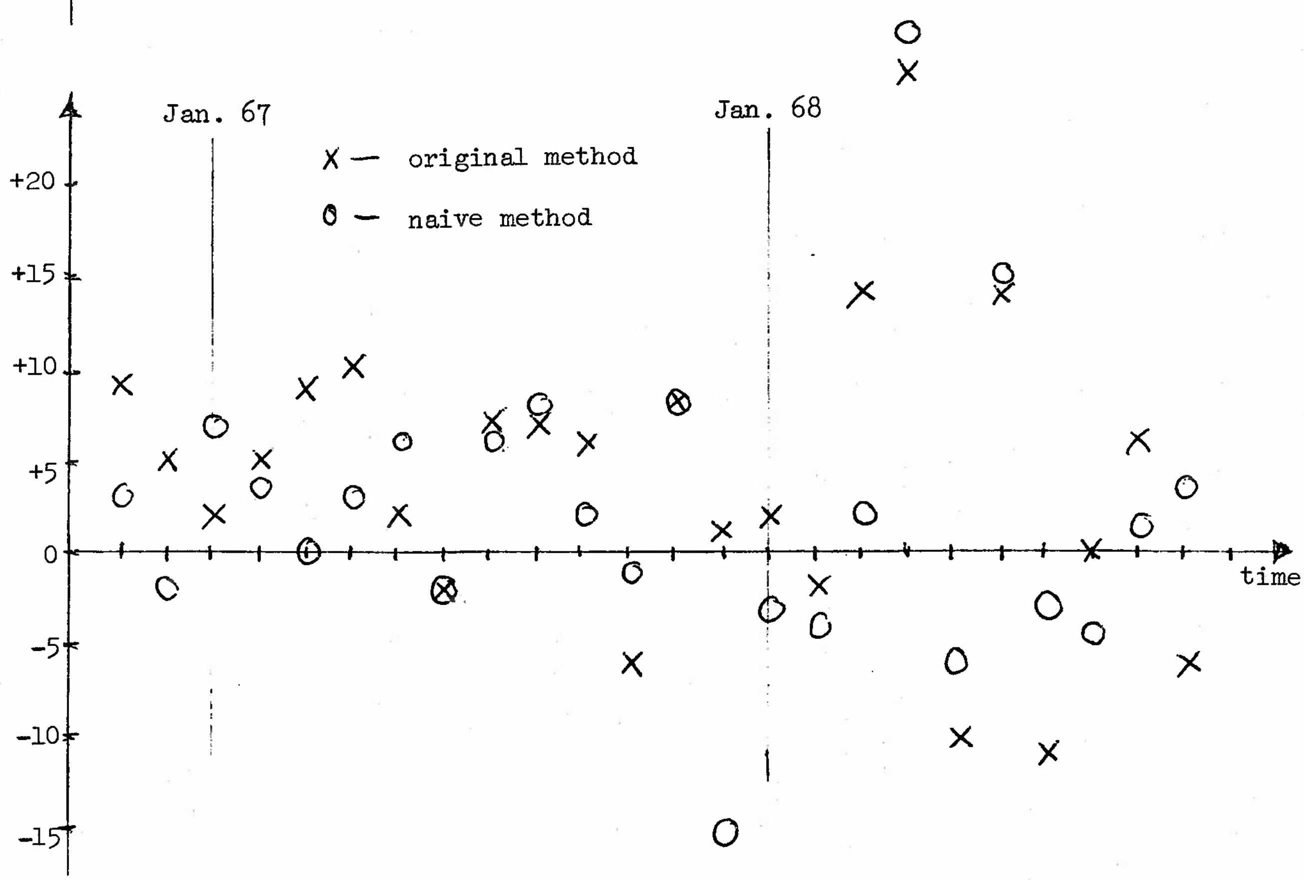

Fig. 2 % Forecast Error in Forecasts of Number of Installations

REFERENCES

Box, G. E. P., and G. M. Jenkins (1970). Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco.

Box, G. E. P., and G. M. Jenkins (1973). Some Comments on a paper by Chatfield and Prothero and on a Review by Kendall. J.R. Statist. Soc. A, 136, Part 3, 337-352.

Chatfield, C., and D. L. Prothero (1973). Box-Jenkins seasonal forecasting: problems in a case study (with Discussion). J.R. Statist. Soc. A., 136, 295-336.

Thompson, H.E., and G. C. Tiao (1971). Analysis of Telephone Data. Bell J. Econ. Manag. Sci., 2, 514-541

Thompson, H. E., and G. C. Tiao (1970). Analysis of Telephone Data. Tech. Report No. 222, Department of Statistics, University of Wisconsin.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.

2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.

3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.

4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.

5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.

6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.

7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.

8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.

9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.

10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.

11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.

12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.

13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.

14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.

15. "Maximum Likelihood Estimation of Parameters of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.

16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time Series," Steffen L. Lauritzen, April, 1974.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER Technical Report No. 16	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) NOTE ON A CASE-STUDY IN BOX-JENKINS SEASONAL FORECASTING OF TIME SERIES	5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Steffen L. Lauritzen	8. CONTRACT OR GRANT NUMBER(s) N00014-67-A-0112-0030
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California 94305	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-034
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 436 Arlington, Virginia 22217	12. REPORT DATE APRIL 11, 1974
	13. NUMBER OF PAGES 7
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Reproduction in whole or in part is permitted for any purpose of the United States Government. Distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

1. time series
2. autoregressive-moving average models
3. forecasting

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

It is shown that in a particular example a straight forward forecasting procedure gives better results than forecasts obtained from a fitted ARIMA model.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)